```
=> d his nofile

     (FILE 'HOME' ENTERED AT 16:03:41 ON 14 FEB 2006)

     FILE 'LREGISTRY' ENTERED AT 16:04:49 ON 14 FEB 2006
L1            0 SEA ABB=ON  YDWRFNAF.Y|YDFRWNAF.Y|YDHFRWAF.Y/SQSFP

     FILE 'REGISTRY' ENTERED AT 16:05:39 ON 14 FEB 2006
L2           21 SEA ABB=ON  YDWRFNAF.Y|YDFRWNAF.Y|YDHFRWAF.Y/SQSFP
                SAVE TEMP L2 GUD394SEQ/A

     FILE 'CAPLUS' ENTERED AT 16:06:22 ON 14 FEB 2006
L3           13 SEA ABB=ON  L2

     FILE 'REGISTRY' ENTERED AT 16:06:26 ON 14 FEB 2006

     FILE 'REGISTRY' ENTERED AT 16:06:40 ON 14 FEB 2006
                D QUE L2
                D RN CN SQL KWIC NTE LC 1-21 L2

     FILE 'CAPLUS, CASREACT, USPATFULL, TOXCENTER' ENTERED AT 16:07:35 ON 14
     FEB 2006
L4           22 SEA ABB=ON  L2
L5           15 DUP REM L4 (7 DUPLICATES REMOVED)
                   ANSWERS '1-13' FROM FILE CAPLUS
                   ANSWERS '14-15' FROM FILE USPATFULL
                D IBIB ED ABS HITRN 1-15

     FILE 'HOME' ENTERED AT 16:08:01 ON 14 FEB 2006

=>
```

## Sequence Family Search of Proteins (/sqsfp)

In the sequence family search, each amino acid in the query has to match either the exact amino acid or a family member equivalent, as shown in the Family Equivalence Table below. The Family Equivalence Table is applied only to each common amino acid in the sequence. Specific uncommon amino acids may be included in the sequence; however, family equivalents only exist for the common amino acids. An amino acid family is based on a conservative substitution of amino acids sharing a similar chemical property. Each common amino acid in the query is converted to its family class members in a search. A match occurs on a query sequence if each amino acid is exactly matched or any of its family members are encountered. For example, the Hydrophobic-Aromatic family consists of the common amino acids F, W, and Y. If the amino acid F is specified within a sequence exact family search, it will match on amino acids F, W, or Y.

### FAMILY EQUIVALENCE TABLE

| Family Class Name | Family Class Members |
|---|---|
| Neutral-Weakly Hydrophobic | Ala (A), Gly (G), Pro (P), Ser (S), Thr (T) |
| Hydrophilic-Acid Amine | Asn (N), Asp (D), Gln (Q), Glu (E) |
| Hydrophilic-Basic | Arg ( R), His (H), Lys (K) |
| Hydrophobic | Ile (I), Met (M), Leu (L), Val (V) |
| Hydrophobic-Aromatic | Phe (F), Trp (W), Tyr (Y) |
| Crosslinking | Cys ( C) |

☐ 1: AAU94398. Reports At1g06100 [Arabid...[gi:53828577]

BLink, Conserved Domains, Links

Features Sequence

```
LOCUS       AAU94398                299 aa            linear   PLN 06-OCT-2004
DEFINITION  At1g06100 [Arabidopsis thaliana].
ACCESSION   AAU94398
VERSION     AAU94398.1  GI:53828577
DBSOURCE    accession BT015835.1
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (residues 1 to 299)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (residues 1 to 299)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-2004) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..299
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="U62926"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     Protein         1..299
                     /product="At1g06100"
     Region          62..263
                     /region_name="Fatty acid desaturase"
                     /note="FA_desaturase"
                     /db_xref="CDD:25523"
     CDS             1..299
                     /coded_by="BT015835.1:1..900"
                     /note="unknown protein"
ORIGIN
        1 msettkddgs sqkksvrkek rayvlrkwtq fdvgrastvg tvhllcllap fnykweafrf
```

```
      61 giilailtnl citfsyhrnl thrsfklpkw leypfaysal lalqgdpldw vsihrfhhqf
     121 tdsdrdphsp iegfwfshvl wifdtdyire kcgrrnnvmd lkqqwfyrfl kktlvlhila
     181 fwtliylwgg lpyltwtvgf ggvigyhgtw lvnsachicg sqawqtndts rnvwwlallt
     241 mgeswhnnhh afetsarhgl ewyqlditwy liwffqalgl atnvklptda qkrkmairr
//
```

☐ 1: AAT47784. Reports At1g06100 [Arabid...[gi:48958463]

BLink, Conserved Domains, Links

Features Sequence

```
LOCUS       AAT47784                299 aa            linear   PLN 19-JUN-2004
DEFINITION  At1g06100 [Arabidopsis thaliana].
ACCESSION   AAT47784
VERSION     AAT47784.1  GI:48958463
DBSOURCE    accession BT014933.1
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (residues 1 to 299)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (residues 1 to 299)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2004) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..299
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="S62926"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     Protein         1..299
                     /product="At1g06100"
     Region          62..263
                     /region_name="Fatty acid desaturase"
                     /note="FA_desaturase"
                     /db_xref="CDD:25523"
     CDS             1..299
                     /coded_by="BT014933.1:1..900"
                     /note="unknown protein"
ORIGIN
        1 msettkddgs sqkksvrkek rayvlrkwtq fdvgrastvg tvhllcllap fnykweafrf
```

```
      61 giilailtnl citfsyhrnl thrsfklpkw leypfaysal lalqgdpldw vsihrfhhqf
     121 tdsdrdphsp iegfwfshvl wifdtdyire kcgrrnnvmd lkqqwfyrfl kktlvlhila
     181 fwtliylwgg lpyltwtvgf ggvigyhgtw lvnsachicg sqawqtndts rnvwwlallt
     241 mgeswhnnhh afetsarhgl ewyqlditwy liwffqalgl atnvklptda qkrkmairr
//
```